Structural basis for regulation of the nucleo-cytoplasmic distribution of Bag6 by TRC35

Jee-Young Mock[a], Yue Xu[b], Yihong Ye[b], and William M. Clemons Jr.[a,1]

[a]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125; and [b]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Bethesda, MD 75534

Edited by Jeffrey L. Brodsky, University of Pittsburgh, Pittsburgh, PA, and accepted by Editorial Board Member F. Ulrich Hartl, September 20, 2017 (received for review February 23, 2017)

The metazoan protein BCL2-associated athanogene cochaperone 6 (Bag6) forms a hetero-trimeric complex with ubiquitin-like 4A and transmembrane domain recognition complex 35 (TRC35). This Bag6 complex is involved in tail-anchored protein targeting and various protein quality-control pathways in the cytosol as well as regulating transcription and histone methylation in the nucleus. Here we present a crystal structure of Bag6 and its cytoplasmic retention factor TRC35, revealing that TRC35 is remarkably conserved throughout the opisthokont lineage except at the C-terminal Bag6-binding groove, which evolved to accommodate Bag6, a unique metazoan factor. While TRC35 and its fungal homolog, guided entry of tail-anchored protein 4 (Get4), utilize a conserved hydrophobic patch to bind their respective partners, Bag6 wraps around TRC35 on the opposite face relative to the Get4–5 interface. We further demonstrate that TRC35 binding is critical not only for occluding the Bag6 nuclear localization sequence from karyopherin α to retain Bag6 in the cytosol but also for preventing TRC35 from succumbing to RNF126-mediated ubiquitylation and degradation. The results provide a mechanism for regulation of Bag6 nuclear localization and the functional integrity of the Bag6 complex in the cytosol.

tail-anchor targeting | X-ray crystallography | tail-anchor recognition complex | proteasome-dependent degradation | GET pathway

Tail-anchored (TA) membrane proteins contain a single, C-terminal, hydrophobic transmembrane-helix domain (TMD) (1). The TMD, which acts as the endoplasmic reticulum (ER) targeting signal, emerges from the ribosome after translation termination necessitating posttranslational targeting and insertion. Transmembrane domain recognition complex (TRC) proteins in mammals and guided entry of tail-anchored proteins (GET) in fungi are the best-characterized molecular players in TA targeting (2). In the TRC pathway, TA proteins are captured by the small glutamine-rich tetratricopeptide-repeat protein A (SGTA) after release from the ribosome then transferred to the ATPase TRC40, a process facilitated by the hetero-trimeric Bag6–TRC35–Ubl4A complex (3, 4). TA binding stimulates hydrolysis of ATP by TRC40 (5) leading to release of the TRC40–TA complex that then localizes to the ER membrane via its interaction with calcium-modulating ligand (CAML) (6). Release of the TA substrate for insertion is facilitated by CAML and tryptophan rich basic protein (WRB) (7). Notably, loss of WRB in mouse cardiomyocytes and hepatocytes results in partial mislocalization of TA proteins (8), suggesting that other pathways, including the signal-recognition particle (SRP) (9), SRP-independent targeting (SND) proteins (10), the Hsc70 family of chaperones (11), and ubiquilins (12), can compensate to target TA in vivo.

Structural information for the mammalian proteins is sparse; however, homology to the equivalent fungal counterparts, Get1–5 and Sgt2, allow for parallels. Bag6 is unique among the mammalian proteins, and its introduction alters the molecular architecture of the central hand-off complex. For TA targeting, the heterotrimeric Bag6 complex is equivalent to the fungal heterotetrameric Get4–Get5 complex (3, 13), but the addition of Bag6 leads to different interactions.

Crystal structures of fungal Get4 revealed 14 α-helices that can be divided into N-terminal domains (NTD) and C-terminal domains (CTD) (Fig. 1*A* and Fig. S1*A*) (14–16). The NTD contains conserved residues (Fig. S1*C*) involved in binding and regulating Get3 (13, 15), while the CTD forms a stable interaction with the Get5 N terminus (Get5-N) (15). The Get5 interface includes a hydrophobic cleft in Get4 formed by a β-hairpin between α13 and α14 that binds the helix in Get5-N (15). Sequence alignment of Get4/TRC35 and Get5/Ubl4A led to predictions that the Get4 β-hairpin and Get5 N-terminal extension are absent in metazoans (Fig. 1 *A* and *B*) (15), which suggested a different structural organization for the metazoan Bag6–TRC35–Ubl4A complex.

In metazoans, Bag6 abrogates the interaction of TRC35 with Ubl4A by separately binding these factors (3, 17) resulting in a heterotrimer. The interface between TRC35 and Bag6 had previously been mapped to the CTD of TRC35 and the region on Bag6 that contains the bipartite nuclear localization sequence (NLS) (3, 18). Overexpression of TRC35 results in Bag6 retention in the cytosol (18), suggesting that TRC35 plays a role in determining the nucleo-cytoplasmic distribution of Bag6. Potential roles for Bag6 in the nucleus are varied and include modulation of p300 acetylation (19, 20), facilitation of histone methylation (21), and regulation of DNA damage signaling-mediated cell death (22). How Bag6 localization is regulated remains unclear.

In this study, we present a crystal structure of a complex that includes TRC35 from *Homo sapiens* (Hs) and the TRC35-binding domain of HsBag6. The structure of TRC35 reveals

Significance

The metazoan protein BCL-2–associated athanogene cochaperone 6 (Bag6) acts as a central hub for several essential cellular processes, including immunoregulation, gene regulation, apoptosis, and proteostasis. These roles are in both the nucleus and the cytosol, but the mechanism by which Bag6 traffics between these compartments remains elusive. Here we present the crystal structure of Bag6 in complex with its cytoplasmic retention factor transmembrane domain recognition complex 35 (TRC35) and suggest a mechanism of regulation for the nucleo-cytoplasmic transport of Bag6.

Author contributions: J.-Y.M., Y.Y., and W.M.C. designed research; J.-Y.M., Y.X., and Y.Y. performed research; J.-Y.M., Y.X., and Y.Y. contributed new reagents/analytic tools; J.-Y.M., Y.Y., and W.M.C. analyzed data; and J.-Y.M. and W.M.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.L.B. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates and structure factors reported in this paper have been deposited in the Protein Data Bank (PDB), https://www.wwpdb.org (PDB ID code 6AU8).

[1]To whom correspondence should be addressed. Email: clemons@caltech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1702940114/-/DCSupplemental.

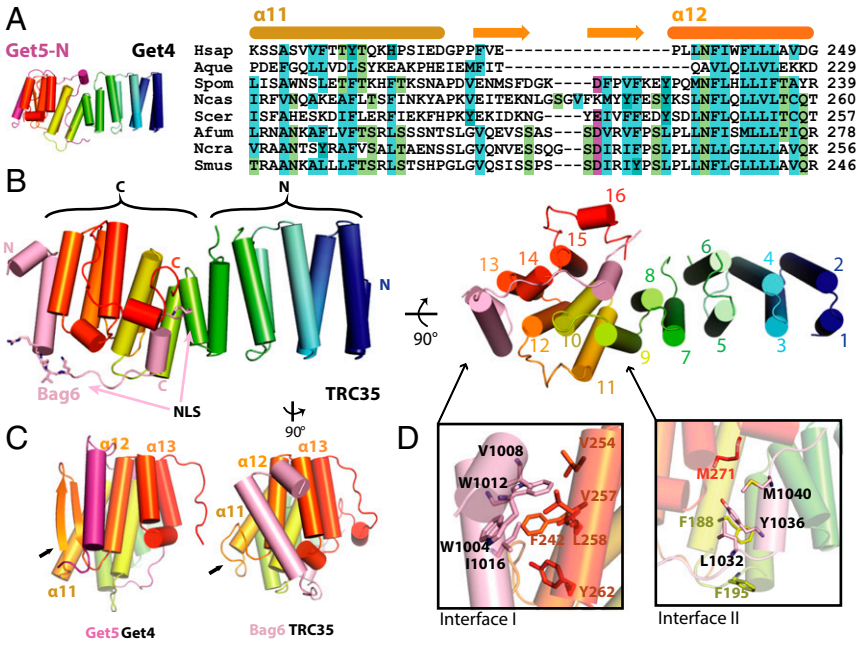

Fig. 1. The structure of the Bag6NLS–TRC35 heterodimer. (*A*) The structure of *Saccharomyces cerevisiae* Get4–Get5N complex [Protein Data Bank (PDB) ID code: 3LKU], Get4 in rainbow and Get5 in magenta. Sequence alignment of TRC35/Get4 from *Homo sapiens*, *Amphimedon queenslandica*, *Schizosaccharomyces pombe*, *Naumovozyma castelli*, *Saccharomyces cerevisiae*, *Aspergillus fumigatus*, *Neurospora crassa*, and *Sphaerulina musiva*. The secondary structure based on fungal Get4s is highlighted above the text. (*B*, *Left*) The overall structure of Bag6–TRC35 heterodimer in cylinder representation with Bag6 in light pink and TRC35 in rainbow. The nuclear localization sequence is highlighted in sticks on Bag6. (*B*, *Right*) A 90° in plane rotated 'bottom' view. The 16 helices of TRC35 are labeled from N to C terminus. (*C*) Comparison of the C-terminal faces of TRC35 and Get4 that bind Bag6 and Get5, respectively. The arrows highlight the significant structural difference in the residues between α11 and α12. (*D*) Zoomed in view of the regions, defined as interface I and II. Hydrophobic residues that are involved in Bag6–TRC35 dimerization are shown as sticks.

the conserved architecture relative to fungal Get4 homologs. Surprisingly, Bag6 utilizes the same conserved pockets on TRC35 that are recognized by Get5 on Get4; yet the extended domain wraps around the opposite face of the α-solenoid. The resulting interaction leads to masking of the bipartite NLS preventing nuclear trafficking of Bag6 by blocking its binding to karyopherin α (KPNA), which is demonstrated experimentally. The interaction also prevents RNF126-mediated degradation of TRC35. These results suggest a mechanism for regulation of Bag6 nucleocytoplasmic distribution.

The Crystal Structure of the TRC35–Bag6–NLS Complex

A complex between TRC35 (residues 23–305), shown to be competent for TA transfer (3), and a minimal TRC35-binding domain of Bag6 that contains the NLS (residues 1,000–1,054) was coexpressed, purified and crystallized. A dataset from a single crystal was collected to 1.8 Å resolution and phases were obtained using single wavelength anomalous dispersion from a rubidium derivative. The final model refined to 1.8–39 Å had good statistics with an R_{free} = 20.1%. Full crystallographic statistics are provided in Table S1. All TRC35 residues in the crystallized construct were visible in the electron density except for S137 and G138 that are disordered in the loop between α6 and α7. For Bag6, residues 1,002–1,043 were resolved with only a few terminal residues ambiguous in the electron density.

TRC35 has the same overall architecture as fungal Get4 (Fig. S1*A*), revealing that the Getfourfold has been conserved across Opisthokonta. Opisthokonta includes metazoan, choanozoan, and fungal lineages (Fig. S2), which share common ancestry as determined by analysis of 16S-like rRNA (23) and protein sequences (24). The first 15 α-helices form an α-solenoid fold that can be divided into N- and C-terminal halves (NTD and CTD) (Fig. 1*B*). Alignment of the NTD between TRC35 and Get4 results in an RMSD = 1.380 Å, while the equivalent alignment in the CTD results in an RMSD = 3.429 Å (Fig. S1*B*). As seen in Get4 (15), there is likely some flexibility between the two halves based on differences across crystal forms. TRC35 and TRC40 are conserved throughout opisthokont evolution and seem to occur as a pair, suggesting the essentiality of the two proteins in the pathway (Fig. S2). Accordingly, residues at the TRC35–TRC40 interface (13) and the TRC35–Bag6 interface in TRC35 are well conserved (Fig. S1 *C* and *D*). As predicted from sequence alignment, the β-hairpin in Get4 is replaced by a shorter loop in TRC35 (Fig. 1*C*, arrows). In fact, sequence alignment of the predicted Get4 proteins of selected opisthokonts revealed that the β-hairpin is unique to the fungal lineage (Fig. S2 and Table S3). The C-terminal α16 is flanked by two extended stretches that cover part of Bag6 (Fig. 1*B*).

The Bag6 NLS region wraps around the TRC35 CTD (Fig. 1*B*, light pink) at an interface that resembles that of Get5–Get4 (Fig. 1*C*). The interaction is stabilized by two hydrophobic interface sites. In interface I, Bag6 α1 and α2 dock into a conserved pocket created by α12 and α13 of TRC35 (Fig. S1*D*) and include W1004, V1008 and W1012 from Bag6 and V254, V257, F242, L258 and Y262 from TRC35 (Fig. 1*D*). In fungal Get4–5, the Get5 N-terminal helix docks into a groove formed by α12, α13, and the β-hairpin of Get4 (Fig. 1*C*). The missing β-hairpin in TRC35 results in the Bag6 helix shifting away from α11 toward α12 and α13 near the bottom of TRC35 (Fig. 1*C*, arrows). The differences in the interface result in rearrangement of α11, α13, and α15 of TRC35 (Fig. 1*C*) relative to those of Get4. Interface II is less extensive involving fewer residues, L1032, Y1036, M1040 of Bag6 and F195, M271 of TRC35 (Fig. 1*D*). While the connecting loops between the two interfaces are well-ordered in both contexts, the Bag6 loop makes fewer contacts to TRC35, buried surface area of 2,841 $Å^2$, than the extensive interactions in the Get5-loop to Get4, buried surface area of 5,554 $Å^2$ (Fig. 1*B*).

The structure reveals how TRC35 masks the Bag6 NLS (Fig. 1*B*). The conserved first basic cluster of the predicted NLS (K_{1024}RVK) is sequestered between interfaces I and II (Fig. 1*B*, sticks). Only the first lysine residue of the second cluster (K_{1043}RRK) is resolved in our model (Fig. 1*B*, sticks). The truncated C-terminal residues of TRC35 in our construct are predicted to be disordered and include conserved multiple negative charges (five Glu and three Asp) (Fig. S3) that would be in position to mask the second NLS basic cluster, K_{1043}RRK, through charge-mediated interactions.

Probing the Bag6–TRC35 Interface

To validate the interfaces observed in the structure, we generated alanine mutants for analysis by yeast two-hybrid (Y2H) analysis. A fragment of Bag6 (residues 951–1,126) was attached

to the GAL4 transcription activating domain and full-length TRC35 was attached to the GAL4 DNA-binding domain. The transformation of both vectors was confirmed by cell growth on SC-Ura-Leu medium (Fig. S4 *A* and *B*). Of the single amino acid substitutions (Fig. 2*A* and Fig. S4 *A* and *B*), only one residue TRC35 (Y262A) at the end of interface I (Fig. 1D) disrupted the Y2H interaction (Fig. S4*A*). The Bag6 mutations W1004A and W1012A localize at interface I and Y1036A localizes at interface II. While the individual mutations did not show a phenotype, the combination of the two mutations (W1004A/Y1036A or W1012A/Y1036A) synthetically disrupted the interaction (Fig. 2*A*). For mutants that failed to grow, expression was confirmed by immunoblotting (Fig. S4*C*). The results confirm that both interfaces are critical for forming a stable complex between Bag6 and TRC35.

We further validated the interaction of TRC35 with full length Bag6 in the context of a mammalian cellular environment. GFP-tagged WT Bag6 or the mutants Bag6(W1004A), Bag6(W1012), Bag6(Y1036A), Bag6(W1004A/Y1036A), and Bag6(W1012A/Y1036A) were coexpressed with FLAG-tagged TRC35 in CRISPR-mediated Bag6-knockout 293T cells (25). Bag6 and associated proteins were captured from cell extracts by immunoprecipitation (IP) using a GFP-Ab. Viewed by immunoblot (Fig. 2*B* and Fig. S4 *D* and *E*), each Bag6 mutation resulted in a reduction of the captured TRC35 relative to wtBag6 (lane 9 vs. lanes 10–12). The double mutants (lanes 13 and 14) also had reduced capture of TRC35, but they unexpectedly appeared to capture more TRC35 than the single mutants. Furthermore, higher molecular weight TRC35 bands appeared that became more pronounced in the double mutants (Fig. 2, asterisk and Fig. S4 *D* and *E*). This will be discussed below. Overall, the general effects of mutations at the Bag6 interface are consistent between the Y2H and the mammalian pulldowns.

Improperly Assembled TRC35 Is Ubiquitylated by RNF126

As noted, IP results (Fig. 2*B* and Fig. S4 *D* and *E*) revealed two interesting observations: (*i*) an increase in TRC35 binding to Bag6 double mutants relative to single mutants (Fig. 2*B* and Fig. S4 *D* and *E*, lanes 11 and 12 vs. lanes 13 and 14) and (*ii*) the appearance of higher molecular weight products of TRC35 captured by Bag6 double mutants (Fig. 2*B* and Fig. S4 *D* and *E*, asterisk). Evidence supports that the stability of TRC35 requires forming a proper complex with Bag6 (17, 22). Given the implication of Bag6 as a chaperone holdase in protein quality control processes such as mislocalized protein degradation (26) and ER-associated protein degradation pathway (18), we postulated that TRC35 mutants that fail to bind Bag6 at its physiological binding site are unstable and become a target for Bag6-mediated degradation. In this case, unassembled TRC35 would bind to the Bag6 substrate-binding site. Because the Bag6 construct used in our Y2H studies lacks the substrate-binding domain, increased binding is not observed in the double mutant experiments. If this model is correct, the higher molecular weight bands should be ubiquitylated TRC35 bound to the substrate-binding site on Bag6.

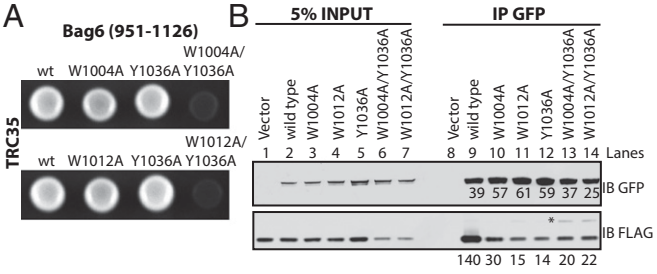

Fig. 2. Validation of the Bag6–TRC35 interface. (*A*) Yeast two-hybrid assay to validate the interface identified in the crystal structure. WT TRC35 conjugated to the DNA binding domain was expressed with WT or mutant Bag6 (951–1126) conjugated to the transcription activating domain. Interaction was determined by ability to grow on adenine-negative medium. (*B*) WT or mutant Bag6•GFP was coexpressed in Bag6$^{-/-}$ 293T cells with TRC35•FLAG and immunoprecipitated using anti-GFP Ab. Amount of TRC35 retrieved by Bag6 was assessed by blotting with anti-FLAG Ab. The position of the higher molecular weight TRC35·FLAG is indicated by an asterisk.

To verify a role for Bag6 in TRC35 ubiquitylation, Bag6 deficient cells were used to coexpress TRC35· holdase FLAG, Bag6 holdase GFP variants and HA holdase ubiquitin (Ub). Proteins bound to Bag6 were first immunoprecipitated with GFP Ab (Fig. S5 *A* and *B*, lanes 2–7). To remove other ubiquitylated Bag6 substrates (18, 27), the samples obtained from the GFP IP were subject to a second round of IP using anti-FLAG beads under denaturing conditions (Fig. S5 *A* and *B*, lanes 9–14). Immunoblotting showed that all Bag6 variants pulled down ubiquitylated proteins, suggesting that the mutations did not affect its substrate-binding activity (Fig. S5 *A* and *B*, lanes 2–7). After the second IP with FLAG Ab, TRC35 associated with wtBag6 carried a small amount of Ub conjugates (Fig. S5 *A* and *B*, lane 9), while those associated with Bag6 variants that disrupted physiological association with TRC35 carried significantly more Ub conjugates (Fig. S5 *A* and *B*, lanes 10–14). Compared with single Bag6 mutations, TRC35 bound to the double mutants had the highest Ub-modified to unmodified TRC35 ratio (Fig. S5 *A* and *B*, lanes 10–12 vs. lanes 13 and 14). Furthermore, the level of TRC35 in cells expressing Bag6 double mutants was consistently lower than in cells expressing wtBag6 (Fig. S5*C*), which could be rescued by treatment with the proteasome inhibitor MG132 (Fig. 3*A*, lane 2 vs. lane 4). Treating cells with MG132 also increased ubiquitylated TRC35 that was captured by either wtBag6 or a Bag6 double mutant (Fig. 3*A*). The accumulation of unmodified TRC35 (Fig. 3*A*, lanes 9 and 10 vs. lanes 11 and 12) is probably due to depletion of free Ub in the cell (28). Collectively, these data support the idea that the TRC35 molecules unable to associate with Bag6 at the NLS domain become targets for Ub-dependent degradation through a client–chaperone interaction with Bag6. This suggests that the ubiquitylated TRC35 associates with the quality control module (QC) of Bag6 (27).

To further ensure that unassembled TRC35 is a Bag6-QC substrate, we investigated the effect of the Ub-ligase RNF126 on TRC35 ubiquitylation. RNF126 is a Bag6-associated Ub-ligase utilized by the Bag6-QC for ubiquitylation of Bag6-associated clients in the cytosol (27). If TRC35 ubiquitylation is mediated by Bag6-QC, knocking down RNF126 in cells expressing Bag6 mutants would result in reduced ubiquitylation and increased stability of TRC35. To test this hypothesis, we examined the effect of siRNA-mediated RNF126 knockdown (KD) on TRC35 ubiquitylation. 293T cells coexpressing HA·Ub, TRC35·FLAG, and either WT or mutant Bag6·GFP (W1004A/Y1036A or W1012A/Y1036A) were treated with siRNA against RNF126 followed by 6-h treatment with MG132. Bag6-associated TRC35 was isolated by two rounds of IP. Immunoblotting was used to analyze the relative ratio between ubiquitylated TRC35 and unmodified TRC35 (Ub–TRC35/TRC35). The ratio in cells expressing both wtTRC35 and wtBag6 without RNF126 was defined as 1. In cells expressing Bag6 mutants, ubiquitylation of TRC35 bound to Bag6 is moderately increased (Fig. 3*B*, lane 9 vs. lane 10); RNF126 KD reduces TRC35 ubiquitylation in cells expressing either wtBag6 or Bag6 double mutants, resulting in approximately two- to threefold reduction in Ub–TRC35/TRC35 (Fig. 3*B*, lanes 9 and 10 vs. lanes 11 and 12). Importantly, like MG132 treatment, RNF126 KD rescued the expression of TRC35 in cells expressing the double mutant (Fig. S6), presumably due to increased stability. Together these results demonstrate that ubiquitylation of unassembled TRC35 is modulated by the QC role of Bag6 and RNF126.

BIOCHEMISTRY

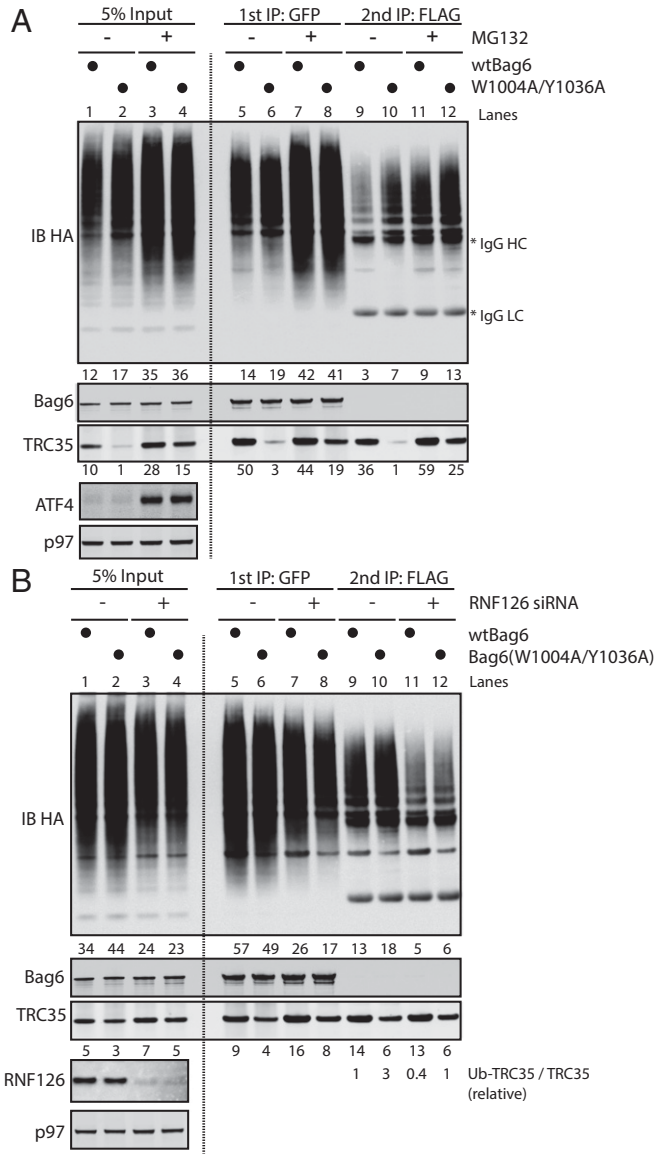

Fig. 3. Ubiquitylation of TRC35 in cells expressing Bag6 mutants is dependent on RNF126. (*A*) Bag6$^{-/-}$ 293T cells cotransfected with plasmids encoding TRC35•FLAG (WT), Bag6·GFP (WT or mutant), and HA•Ub and treated with 10 μM MG132. The cell extracts were subject to two rounds of IP with anti-GFP Ab then with anti-FLAG Ab in denaturing conditions. Ubiquitylation was assessed by immunoblotting with HA Ab. (*B*) Bag6$^{-/-}$ 293T cells cotransfected with plasmids encoding TRC35•FLAG (WT), Bag6•GFP (WT or mutants), and HA·Ub. The cells were simultaneously treated with 10 μM MG132 and siRNAs against RNF126. The cell extracts were subject to two rounds of denaturing IP with anti-GFP Ab and then with anti-FLAG Ab. TRC35 ubiquitylation was assessed by immunoblotting with HA Ab.

TRC35 Masks the Bipartite Nuclear Localization Sequence of Bag6 and Retains Bag6 in the Cytosol

To investigate the Bag6–TRC35 interface in the context of Bag6 localization, the mutants identified from Y2H and IP were used for localization studies. Overexpression of TRC35 has been shown to retain Bag6 in the cytosol (18), suggesting that TRC35 binding is required for cytosolic localization of Bag6. Because the mutations identified in Y2H experiments specifically prevent TRC35 binding at its native binding site, exogenously expressed Bag6 mutants defective in TRC35 binding would localize primarily in the nucleus regardless of TRC35 expression. To test this hypothesis, WT or mutant Bag6•GFP were expressed in Cos7 cells with or without TRC35•FLAG. The localization of Bag6 and TRC35 was examined by immunofluorescence.

As expected (18), overexpressed wtBag6 and various Bag6 mutants were localized to the nucleus (Fig. 4*A*). This is likely due to excess Bag6 that cannot be retained in the cytosol by endogenous TRC35. When TRC35 is coexpressed with wtBag6, the increased cytosolic pool of TRC35 captures wtBag6 and both proteins stain primarily in the cytosol (Fig. 4*A*). In accordance with the IP results (Fig. 2*B*), introduction of a single mutation—W1004A, W1012A, or Y1036A—results in some Bag6 localization to the nucleus (Fig. 4 *B–D* and Table S3). Mutations at both interface I (W1004A or W1012A) and interface II (Y1036A) further disrupt the binding between Bag6 and TRC35 (Fig. 2), and Bag6 localizes primarily to the nucleus (Fig. 4 E and F and Table S3). These results unequivocally establish TRC35 as a cytoplasmic retention factor of Bag6.

TRC35 Has Higher Affinity for Bag6 than Karyopherin-α 2

All molecules destined for the nucleus must move through the nuclear pore complex, which spans the nuclear envelope and facilitates nucleo-cytoplasmic traffic (29). Macromolecules larger than ~40 kDa cannot freely diffuse through and require carrier proteins, such as the karyopherin-α (KPNA) family of transporters (30). The two basic clusters, R_{1024}KVK and K_{1043}RRK (Fig. S7*A*), in Bag6 are thought to act as a bipartite NLS by specifically recognizing the acidic substrate-binding surface of karyopherins (31). In HeLa cells, the K_{1045}R to S_{1045}L mutation has been shown to abrogate nuclear localization of Bag6 (31). Therefore, TRC35 and KPNA likely bind Bag6 at the fragment that contains the NLS.

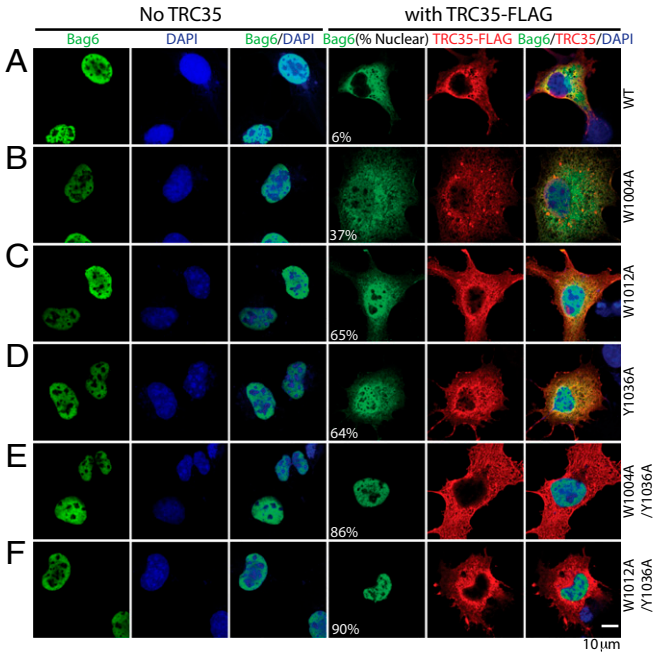

Fig. 4. Bag6 mutations at the TRC35 binding site results in nuclear localization of Bag6. (*A–F*) Representative images of Cos7 cells expressing WT TRC35·FLAG and WT or mutant Bag6·GFP. Cos7 cells were transfected either with Bag6·GFP (WT or mutant) expressing plasmid alone or cotransfected with TRC35•FLAG (WT) expressing plasmid. Cells were stained with anti-GFP (green) and/or anti-FLAG (red) Abs. DNA was stained with DAPI where indicated (blue). In the Bag6 column, the percent of cells that contain nuclear Bag6 is noted in parenthesis. For each Bag6 variant, ~100 cells were counted and categorized as having nuclear or cytosolic Bag6.

To confirm that the Bag6-NLS is a KPNA-binding site and to define the residues involved in KPNA binding, 131 C-terminal residues of Bag6 (Bag6C131) (32) in complex with hexahistidine-tagged full-length Ubl4A (Bag6C131–6xHis·Ubl4A or Bag6–Ubl4A) was purified from *E. coli*. KPNA2 has been shown to interact with Bag6 (33) and could be stably purified. Mutations previously shown to abrogate Bag6 nuclear localization (31) were introduced at either the first basic cluster (1024SL) or the second basic cluster (1045SL) (Fig. S7*A*). The purified Bag6–Ubl4A variants were incubated with either GST·TRC35 or MBP·KPNA2 lacking the importin β binding domain (58–529), and the resulting complexes were isolated using Ni-NTA beads. WT, 1024SL, and 1045SL Bag6-Ubl4A pulled down GST·TRC35 with similar efficiency (Fig. S7*B*, lanes 5–7), demonstrating that these residues are not involved in binding TRC35. MBP·KPNA2 formed a complex with Bag6 that can be visualized in a pulldown (Fig. S7*B*, lane 9). Surprisingly, only mutation of the second basic cluster led to disruption of the interaction between Bag6–Ubl4A and MBP·KPNA2 (Fig. S7*B*, lane 10 vs. lane 11). These results show that the residues required for binding KPNA2 are distinct from those required for binding TRC35. Moreover, unlike what had been previously suggested (31), the Bag6 NLS is monopartite with the second basic cluster both necessary and sufficient for binding KPNA2. The overlap of the two binding sites provides a mechanism for Bag6 retention in the cytoplasm.

If binding of TRC35 to Bag6 prevents KPNA-mediated translocation, only TRC35-free Bag6 should be able to bind KPNA and be translocated to the nucleus. One way this could be achieved is if KPNAs have a higher affinity for Bag6 than TRC35, up-regulation of KPNA expression would lead to displacement of TRC35 from Bag6 and formation of a Bag6–KPNA complex. To test this hypothesis, we compared the relative binding affinities of TRC35 and KPNA2 to Bag6 using an exchange assay. Bag6–Ubl4A complexes with GST•TRC35 were generated and bound to glutathione affinity resin. After washing, increasing amounts of KPNA2 were added. The ability of KPNA2 to displace Bag6–Ubl4A from TRC35 was determined by the amount of Bag6–Ubl4A eluted from the resin after incubation. In this case, even at the highest concentration tested (twofold molar excess), there was no significant displacement of Bag6 from TRC35 by KPNA2 (Fig. 5, lanes 4–7). When starting with MBP·KPNA2–Bag6–Ubl4A on amylose beads, adding excess TRC35 resulted in the dissociation of the KPNA2–Bag6–Ubl4A complex (Fig. 5, lanes 11–14). These results highlight the stability of the TRC35–Bag6 complex and argue against the ability of KPNA regulation as a means for modulating the nuclear pool of Bag6.

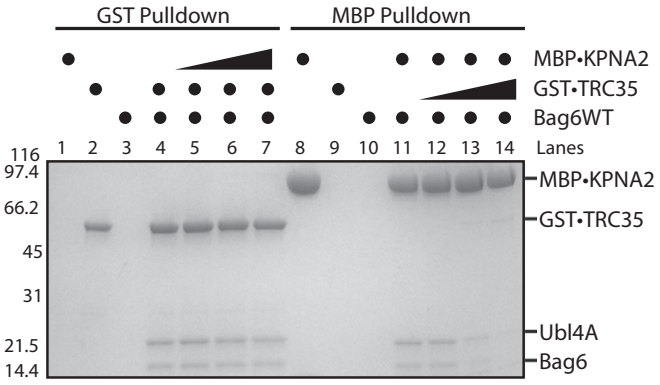

Fig. 5. TRC35 binding precludes karyopherin α binding to Bag6. Recombinantly purified Bag6C131–6xHis·Ubl4A (500 pmol) was first incubated with either GST·TRC35 or MBP·KPNA2 (500 pmol). The resulting GST·TRC35–Bag6–Ubl4A complex was incubated with glutathione resin beads, and increasing amounts of MBP·KPNA2 (50, 250, or 1,000 pmol) was added. The ability of MBP·KPNA2 to displace Bag6–Ubl4A from GST·TRC35 was examined by eluting the GST·TRC35 and bound Bag6–Ubl4A from the glutathione resin. The opposite experiment was also carried out using amylose resin beads.

Discussion

Bag6 is a critical scaffolding factor that forms a stable complex with TRC35 and Ubl4A and is involved in protein targeting, protein quality control, and transcription regulation. While there is ample evidence that Bag6 plays both cytosolic and nuclear roles, it is unclear how the localization of Bag6 is regulated. Here we report the structure of the Bag6 NLS region bound to its cytosolic retention factor TRC35, and suggest a mechanism for the regulation of Bag6 localization.

The remarkable conservation of TRC35 both in sequence (Fig. S3) and at the structural level (Fig. S1) from yeast to human can be attributed to the role of TRC35 as a hub of protein–protein interactions in the TRC pathway. The most important interaction, based on its conservation, is between TRC35 and TRC40 (Fig. S2). This interaction is critical for the regulation of TA protein transfer (3, 13) and, as shown in fungal Get4, for release of Get3/TRC40 from the ER membrane to restore the cytosolic pool of complex ready for TA protein transfer (34). The importance of linking TRC35 homologs to Ubl4A homologs is sustained in humans despite the differences in structural features (15). In metazoans, the component Bag6 acts as a scaffold to link TRC35 to Ubl4A. The Bag6–TRC35 structure reveals that TRC35 binds this binding partner utilizing conserved patches in a distinct fashion (Fig. 1). This led to the expansion of the TRC35 protein–protein interaction network while maintaining its crucial interaction with TRC40.

Our results also show that TRC35 acts as an intermolecular mask to Bag6 NLS in a role similarly performed in other pathways. Examples of other cytosolic retention factor pairs include IκB and NF-κB (35) and BRAP2 that retains HMG20A (36). Unlike other cytoplasmic retention factors, which have only been shown to bind their target NLS-containing proteins for occlusion of the NLS, TRC35 also plays a distinct role in the cytoplasmic TA targeting and protein quality control when in complex with Bag6. This dual functionality may be evolutionarily conserved. In yeast, fungal Get4 binds the NTD of Get5, which appears to contain a functional NLS that directs Get5 to the nucleus during a Get5-mediated stress response (37).

These results allow speculation of possible regulatory mechanism for Bag6 nuclear localization. First, disrupting the Bag6–TRC35 interface by introducing alanine mutations led to ubiquitylation of TRC35 (Fig. S5), which is mediated by RNF126. Knocking down Bag6 has been shown to reduce levels of TRC35 in HeLa cells (22), suggesting that Bag6 is required for TRC35 stability in vivo. We also showed that KPNA2 has a lower affinity for Bag6 than TRC35 (Fig. 5), which suggests that for Bag6 to bind KPNA and translocate into the nucleus, it needs to be free of TRC35. The most likely explanation is that cells regulate Bag6 localization by modulating Bag6 levels or decreasing TRC35 levels. In humans, the Bag6 rs3117582 single nucleotide polymorphism found in the promoter region of Bag6 likely affects expression levels and is associated with higher incidence of lung cancer (38–40) and osteoarthritis (41). The localization could also be pretranslationally regulated with differential splicing. In brain and breast tissue, for instance, Bag6 isoforms that lack the NLS are expressed at higher levels than isoforms with the NLS (42).

Our findings establish TRC35 as a universally conserved protein in the opisthokont lineage in both structure and function. These findings support the model in which TRC35 is an important hub of a protein–protein interaction network with its interaction with TRC40 at the core. Higher eukaryotes expanded on this network by the addition of Bag6. Our biochemical studies unequivocally establish TRC35 as a cytosolic retention factor of Bag6 and suggest that Bag6 needs to be in excess for nuclear

localization. However, how different cell types regulate Bag6 localization, whether the regulation can be mediated by specific stress, and the biological implications of differential distribution of Bag6 are important questions for future studies.

Materials and Methods

Detailed descriptions of experiments are provided in *SI Materials and Methods*. Briefly, for crystallization and in vitro binding assays human genes were subcloned and expressed in *Escherichia coli*. Proteins were purified using affinity chromatography, anion-exchange, and size-exclusion chromatography. Truncated fragments of Bag6 and TRC35 were crystallized, and the structure was determined using molecular-replacement single-wavelength anomalous dispersion. The genomes were surveyed using MEME suite (43). The IP assays were carried out using 293T cell extracts. A combination of three siRNAs against RNF126 was added to 293T cells. MG132 (10 μM) was added to inhibit the proteasome. Cos7 cells were used for localization assays.

ACKNOWLEDGMENTS. We thank Daniel Lin and Jens Kaiser for help with data processing, Catherine Day for technical support, members of the W.M.C. laboratory for support and useful discussions, Gordon and Betty Moore for support of the Molecular Observatory at California Institute of Technology, and the staff at the Stanford Synchrotron Radiation Lightsource for assistance with synchrotron data collection. W.M.C. is supported by NIH Grant R01GM097572.

1. Kalbfleisch T, Cambon A, Wattenberg BW (2007) A bioinformatics approach to identifying tail-anchored proteins in the human genome. *Traffic* 8:1687–1694.
2. Stefanovic S, Hegde RS (2007) Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell* 128:1147–1159.
3. Mock JY, et al. (2015) Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain. *Proc Natl Acad Sci USA* 112:106–111.
4. Schuldiner M, et al. (2008) The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134:634–645.
5. Rome ME, Rao M, Clemons WM, Shan SO (2013) Precise timing of ATPase activation drives targeting of tail-anchored proteins. *Proc Natl Acad Sci USA* 110:7666–7671.
6. Yamamoto Y, Sakisaka T (2012) Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. *Mol Cell* 48:387–397.
7. Vilardi F, Lorenz H, Dobberstein B (2011) WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. *J Cell Sci* 124: 1301–1307.
8. Rivera-Monroy J, et al. (2016) Mice lacking WRB reveal differential biogenesis requirements of tail-anchored proteins in vivo. *Sci Rep* 6:39464.
9. Abell BM, Pool MR, Schlenker O, Sinning I, High S (2004) Signal recognition particle mediates post-translational targeting in eukaryotes. *EMBO J* 23:2755–2764.
10. Aviram N, et al. (2016) The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* 540:134–138.
11. Abell BM, Rabu C, Leznicki P, Young JC, High S (2007) Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. *J Cell Sci* 120: 1743–1751.
12. Itakura E, et al. (2016) Ubiquilins chaperone and triage mitochondrial membrane proteins for degradation. *Mol Cell* 63:21–33.
13. Gristick HB, et al. (2014) Crystal structure of ATP-bound Get3-Get4-Get5 complex reveals regulation of Get3 by Get4. *Nat Struct Mol Biol* 21:437–442.
14. Bozkurt G, et al. (2010) The structure of Get4 reveals an alpha-solenoid fold adapted for multiple interactions in tail-anchored protein biogenesis. *FEBS Lett* 584: 1509–1514.
15. Chartron JW, Suloway CJ, Zaslaver M, Clemons WM, Jr (2010) Structural characterization of the Get4/Get5 complex and its interaction with Get3. *Proc Natl Acad Sci USA* 107:12127–12132.
16. Chang YW, et al. (2010) Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1. *J Biol Chem* 285:9962–9970.
17. Mariappan M, et al. (2010) A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* 466:1120–1124.
18. Wang Q, et al. (2011) A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. *Mol Cell* 42:758–770.
19. Sebti S, et al. (2014) BAT3 modulates p300-dependent acetylation of p53 and autophagy-related protein 7 (ATG7) during autophagy. *Proc Natl Acad Sci USA* 111: 4115–4120.
20. Sasaki T, et al. (2007) HLA-B-associated transcript 3 (Bat3)/Scythe is essential for p300-mediated acetylation of p53. *Genes Dev* 21:848–861.
21. Wakeman TP, Wang Q, Feng J, Wang XF (2012) Bat3 facilitates H3K79 dimethylation by DOT1L and promotes DNA damage-induced 53BP1 foci at G1/G2 cell-cycle phases. *EMBO J* 31:2169–2181.
22. Krenciute G, et al. (2013) Nuclear BAG6-UBL4A-GET4 complex mediates DNA damage signaling and cell death. *J Biol Chem* 288:20547–20557.
23. Wainright PO, Hinkle G, Sogin ML, Stickel SK (1993) Monophyletic origins of the metazoa: An evolutionary link with fungi. *Science* 260:340–342.
24. Baldauf SL, Palmer JD (1993) Animals and fungi are each other's closest relatives: Congruent evidence from multiple proteins. *Proc Natl Acad Sci USA* 90:11558–11562.
25. Xu Y, Anderson DE, Ye Y (2016) The HECT domain ubiquitin ligase HUWE1 targets unassembled soluble proteins for degradation. *Cell Discov* 2:16040.
26. Hessa T, et al. (2011) Protein targeting and degradation are coupled for elimination of mislocalized proteins. *Nature* 475:394–397.
27. Rodrigo-Brenni MC, Gutierrez E, Hegde RS (2014) Cytosolic quality control of mislocalized proteins requires RNF126 recruitment to Bag6. *Mol Cell* 55:227–237.
28. Melikova MS, Kondratov KA, Kornilova ES (2006) Two different stages of epidermal growth factor (EGF) receptor endocytosis are sensitive to free ubiquitin depletion produced by proteasome inhibitor MG132. *Cell Biol Int* 30:31–43.
29. Hoelz A, Debler EW, Blobel G (2011) The structure of the nuclear pore complex. *Annu Rev Biochem* 80:613–643.
30. Pumroy RA, Cingolani G (2015) Diversification of importin-α isoforms in cellular trafficking and disease states. *Biochem J* 466:13–28.
31. Manchen ST, Hubberstey AV (2001) Human Scythe contains a functional nuclear localization sequence and remains in the nucleus during staurosporine-induced apoptosis. *Biochem Biophys Res Commun* 287:1075–1082.
32. Wu YH, Shih SF, Lin JY (2004) Ricin triggers apoptotic morphological changes through caspase-3 cleavage of BAT3. *J Biol Chem* 279:19264–19275.
33. Rouillard AD, et al. (2016) The harmonizome: A collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database (Oxford)* 2016:baw100.
34. Rome ME, Chio US, Rao M, Gristick H, Shan SO (2014) Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway. *Proc Natl Acad Sci USA* 111:E4929–E4935.
35. Beg AA, et al. (1992) I kappa B interacts with the nuclear localization sequences of the subunits of NF-kappa B: A mechanism for cytoplasmic retention. *Genes Dev* 6: 1899–1913.
36. Davies RG, Wagstaff KM, McLaughlin EA, Loveland KL, Jans DA (2013) The BRCA1-binding protein BRAP2 can act as a cytoplasmic retention factor for nuclear and nuclear envelope-localizing testicular proteins. *Biochim Biophys Acta* 1833: 3436–3444.
37. Arhzaouy K, Ramezani-Rad M (2012) Nuclear import of UBL-domain protein Mdy2 is required for heat-induced stress response in Saccharomyces cerevisiae. *PLoS One* 7: e52956.
38. Etokebe GE, et al. (2015) Association of the FAM46A gene VNTRs and BAG6 rs3117582 SNP with non small cell lung cancer (NSCLC) in Croatian and Norwegian populations. *PLoS One* 10:e0122651.
39. Zhao J, Wang H, Hu W, Jin Y (2014) Effect of HLA-B-associated transcript 3 polymorphisms on lung cancer risk: A meta-analysis. *Med Sci Monit* 20:2461–2465.
40. Chen J, Zang YS, Xiu Q (2014) BAT3 rs1052486 and rs3117582 polymorphisms are associated with lung cancer risk: A meta-analysis. *Tumour Biol* 35:9855–9858.
41. Etokebe GE, et al. (2015) Susceptibility to large-joint osteoarthritis (hip and knee) is associated with BAG6 rs3117582 SNP and the VNTR polymorphism in the second exon of the FAM46A gene on chromosome 6. *J Orthop Res* 33:56–62.
42. Luce MJ, Akpawu AA, Tucunduva DC, Mason S, Scott MS (2016) Extent of pre-translational regulation for the control of nucleocytoplasmic protein localization. *BMC Genomics* 17:472.
43. Bailey TL, et al. (2009) MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res* 37:W202–W208.
44. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
45. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
46. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
47. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W (2003) Generation, representation and flow of phase information in structure determination: Recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* 59: 2023–2030.
48. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* 63:32–41.
49. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
50. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
51. Eme L, Sharpe SC, Brown MW, Roger AJ (2014) On the age of eukaryotes: Evaluating evidence from fossils and molecular clocks. *Cold Spring Harb Perspect Biol* 6:a016139.
52. Bailey TL, Gribskov M (1998) Combining evidence using p-values: Application to sequence homology searches. *Bioinformatics* 14:48–54.
53. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
54. Gietz RD, Schiestl RH (2007) Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2:38–41.
55. Schrödinger, LLC (2015) *The PyMOL Molecular Graphics System* (Schrödinger, LLC, New York), Version 2.0.
56. Pettersen EF, et al. (2004) UCSF Chimera–A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
57. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30:772–780.
58. Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2:28–36.